What is claimed is:

1. A computer-implemented method for analyzing gene expression wherein the method comprises the steps of:

(a) compiling data comprising a plurality of measured gene expression signals into a form suitable for computer-based analysis; and

(b) analyzing the compiled data using iterative independent component analysis (ICA), wherein the analyzing comprises identifying an optimum number of independent clusters into which the data may be grouped.

2. The computer-implemented method of claim 1, where the plurality of measured signals comprise hybridization data for a plurality of known gene sequences.

3. The computer-implemented method of claim 1, wherein the number of independent clusters identified by iterative ICA is correlated to the pattern of gene expression for at least one cell type.

4. The computer implemented method of claim 1, wherein the method correlates at least one of the measured signals to an underlying source generating the signal and experimental noise.

5. The computer implemented method of claim 1, wherein the method correlates at least one of the measured signals, x, to the underlying sources, s, generating the signal, and experimental noise, n, as a function of time, t, such that: $x(t) = f(s(t)) + n(t))$.

6. The computer implemented method of claim 1, wherein the iterative ICA comprises an algorithm that yields three matrices: (i) a set of basis functions ($s_1(t)$, $s_2(t)$ . . . $s_M(t)$,) for each of the genes under study; (ii) a mixing matrix, and (iii) a separating matrix, wherein the separating matrix is the inverse of the mixing matrix.

7. The computer implemented method of claim 1, wherein the data is transformed to describe the logarithm of the ratio of two signals, $y_i(t)= \log_2(R_i(t)/G_i(t))$, where R and G represent the measured signal for gene *i* measured under two different experimental conditions, over time, t.

8. The computer-implemented method of claim 1, wherein the plurality of measured signals comprise a plurality of known DNA sequences hybridized to mRNA isolated from the at least one cell type.

9. The computer-implemented method of claim 2, wherein the plurality of known DNA sequences are arranged to form a solid-state array.

10. The computer-implemented method of claim 1, wherein the number of independent clusters into which the data may be grouped, n, is estimated as a preset number, n_0.

11. The computer-implemented method of claim 10, wherein the data are evaluated by sequentially increasing the number of independent clusters from a minimum number, n_0, and determining the relative fit of the data using n_0 as compared to the new value of n.

12. The computer-implemented method of claim 11, wherein the number of clusters are increased incrementally by one for each evaluation, such that the number of independent clusters into which the data may be grouped increases at each step from n_0, to n_0+1, to n_0+2, until the optimum number of groups (n_{opt}) is determined.

13. The computer-implemented method of claim 1, further comprising dynamic ICA such that the resulting model at least in part describes how the system changes over time.

14. The computer-implemented method of claim 1, further comprising hierarchical ICA such that the complexity of the computational analysis is reduced as the analysis proceeds by removing as inputs data that has been described at earlier stages of the analysis from the set of data points still remaining to be characterized.

15. The computer-implemented method of claim 1, further comprising normalizing the variance of the data.

16. The computer-implemented method of claim 1, further comprising determining if there is a cross-correlation between at least two measured signals within a cluster group, wherein a positive cross-correlation comprises the situation in which the expression of one gene in the group is statistically correlated with the expression of a second gene in the same group.

17. The computer-implemented method of claim 16, wherein the relationship between genes within a group is expressed as a mathematical model describing relative levels of gene expression.

18. The method of claim 17, wherein the model comprises the expression $y_i = f(y_1, y_2, \ldots y_N, u_1, \ldots . u_M) + e$, where $y_1, y_2 \ldots y_N$, is the expression of genes 1, 2, . . . N; $u_1, u_2 \ldots u_M$, corresponds to environmental factors 1, 2 . . . M, and experimental noise is defined by e.

19. The computer-implemented method of claim 17, further comprising the step of removing statistically weak links from the model.

20. A computer-implemented method for analyzing gene expression comprising:

(a) compiling data comprising a plurality of measured signals into a form suitable for computer-based analysis;

(b) applying iterative independent component analysis to cluster the data into an optimal number, n, of independent groups, wherein genes in one independent group comprise expression profiles that are substantially independent of the expression profiles for genes in the other groups; and

(c) determining if there is a cross-correlation between at least two genes within a cluster group, wherein a positive cross-correlation comprises the situation in which the expression of one gene in the group is statistically correlated with the expression of a second gene in the same group.

21. An isolated nucleic acid molecule comprising a sequence comprising a gene expression profile and/or gene function as identified using iterative independent component analysis.